Multi-scale Modelling of Natural Composites Using Thermodynamics-Based Artificial Neural Networks and Dimensionality Reduction Techniques

Giovanni Piunno[1](✉), Ioannis Stefanou[2], and Cristina Jommi[1]

[1] Politecnico di Milano, Piazza Leonardo da Vinci 23, 20129 Milan, Italy
{giovanni.piunno,cristina.jommi}@polimi.it
[2] Nantes Université, Ecole Centrale Nantes, CNRS, Institut de Recherche en Génie Civil et Mécanique (GeM), UMR 6183, 44000 Nantes, France
ioannis.stefanou@ec-nantes.fr

Abstract. Modelling natural composites, as the majority of real geomaterials, requires facing their intrinsic multiscale nature. This allows to consider multi-physics coupling occurring at the microscale, then reflected onto the macroscopic behavior. Geotechnics is constantly requiring reliable constitutive models of natural composites to solve large-scale engineering problems accurately and efficiently. This need motivates the contribution. To capture in detail the macroscopic effects of microscopic processes, many authors have developed multi-scale numerical schemes. A common drawback of such methods is the prohibitive computational cost. Recently, Machine Learning based approaches have raised as promising alternatives to traditional methods. Artificial Neural Networks – ANNs – have been used to predict the constitutive behaviour of complex, heterogeneous materials, with reduced calculation costs. However, a major weakness of ANN is the lack of a rigorous framework based on principles of physics. This often implies a limited capability to extrapolate values ranging outside the training set and the need of large, high-quality datasets, on which performing the training. This work focuses on the use of Thermodynamics-based Artificial Neural Networks – TANN – to predict the constitutive behaviour of natural composites. Dimensionality reduction techniques – DRTs – are used to embed information of microscopic processes into a lower dimensional manifold. The obtained set of variables is used to characterize the state of the material at the macroscopic scale. Entanglement of DRTs with TANN allows to reproduce the complex nonlinear material response with reduced computational costs and guarantying thermodynamic admissibility. To demonstrate the method capabilities an application to a heterogeneous material model is presented.

Keywords: TANN · Thermodynamics · Dimensionality reduction · Multiscale modelling · Composites

A. Ferrari et al. (Eds.): CNRIG 2023, SSGG, pp. 365–372, 2023.
https://doi.org/10.1007/978-3-031-34761-0_45

1 Introduction

Multiscale simulation and homogenization have become crucial tools for modeling complex materials, but require substantial computational resources, rendering their use in industry and engineering practice difficult.

The growth of Artificial intelligence and more advanced computational resources has led to an increasing use of machine learning-based methodologies to improve the efficiency and accuracy of multiscale simulations. Machine learning (ML) algorithms often use Dimensionality Reduction Techniques (DRTs) as an essential step in data pre-processing. There are many works in literature that make use of ML tools for the multiscale modeling of composite materials. [1] focuses on the use of ML techniques to develop multiscale models for multi-permeability porous materials, while [2] focuses on the use of Graph-Informed Neural Networks for solving general multiscale physics problems, among many others (e.g., see [3]).

A limited number of works rely on physics-aware ANNs to speed up multiscale simulations, resulting in black boxes whose results are difficult to justify from a physical standpoint. In this paper, we use Thermodynamics-based Artificial Neural Network (TANN), see [4–6], for the multiscale modeling of micro-structured materials. TANN is coupled with DRTs, the latter applied to microscopic information gathered from numerical simulations to identify a set of Internal State Variables to use at the macroscale. We provide a comparison of several DRTs, including POD, ICA, kernel PCA and autoencoders, and discuss the advantages and disadvantages of each approach for the goal at hand. We propose an application using Drucker-Prager hardening elastic-plastic model with cap, for a heterogeneous model of voxels with spatially correlated constitutive parameters, that mimic a natural composite.

The work is articulated starting from a theoretical section summarizing the main aspects of the employed methods (Sect. 2). In Sect. 3, the realization of the material model and of the numerical database is described. Section 4 details the results obtained. The article ends with conclusions on the obtained results.

2 Theoretical Framework

2.1 Dimensionality Reduction Techniques

The Proper Orthogonal Decomposition (POD) [7] and Principal Component Analysis (PCA) [8] are linear techniques that are used to extract important features or patterns in a dataset. POD, also known as the Karhunen-Loève decomposition, decomposes a dataset into a set of orthogonal modes that represent the most significant features. PCA, on the other hand, finds the directions in a dataset that account for the most variance. The POD modes can be derived from the principal components, and the hierarchy of the modes is determined by the singular values. The method is simply expressed by the following modal decomposition formula:

$$X = \tilde{U}\tilde{S}\tilde{V}^* \rightarrow Z = \tilde{U}^* X \tag{1}$$

Z is the reduced dataset, containing the POD coefficients, projections of X onto the POD modes.

Independent Component Analysis (ICA), see [9], is a technique that aims to identify the underlying independent sources within a multivariate signal by assuming that the data is a linear combination of non-Gaussian sources. Mathematically, ICA can be represented as an optimization problem in which the objective is to uncover a linear transformation (matrix W) that maximizes the non-Gaussianity of the transformed data. Namely, $Z = WX$.

In KPCA [10], a kernel function is used to map the data into a higher dimensional feature space in which they become linearly separable. The first step in KPCA is to compute the kernel matrix, which is a symmetric matrix whose entries are given by the kernel function applied to all pairs of data points. The kernel matrix is then decomposed into its eigenvectors and eigenvalues. The eigenvectors with the largest eigenvalues are selected as the principal components. Finally, the original data is projected onto the principal components, resulting in a lower-dimensional representation of the data.

Autoencoders (AE) are nonlinear dimensionality reduction neural networks, see [8]. AE are unsupervised learning algorithms that map inputs to intrinsic representations and then back to themselves. Given an input $\mathcal{I} \in \mathbb{R}^n$, AE learn an intrinsic representation $\mathcal{R} \in \mathbb{R}^l$, $l \ll n$, which is mapped back into $\mathcal{I}^* \in \mathbb{R}^n$, imposing $\mathcal{I}^* = \mathcal{I}$. The parametrization is implemented by two functions: an encoder, $NN_E : \mathbb{R}^n \to \mathbb{R}^l$, and a decoder, $NN_D : \mathbb{R}^l \to \mathbb{R}^n$.

2.2 Random Field Generation Algorithms

There are numerous published methods for generating realizations of stationary homogeneous spatially correlated random fields. The matrix decomposition method, as described in [11], requires the definition of a discrete set of N_p points at which the random field will be sampled and then to create a covariance matrix, C, quantifying the correlation between all sampling points. With the Cholesky decomposition, the lower triangular matrix of C, L, is obtained. Then, a vector of correlated random variables, $\mathcal{Y}$, is computed by generating a vector of uncorrelated random numbers, $\mathcal{X}$, from a unit normal distribution, calculating $\mathcal{Y} = L\mathcal{X}$. Cholesky decomposition is an exact method, so the simulated Gaussian field follows an exact multivariate Gaussian distribution.

2.3 Thermodynamics-Based Artificial Neural Networks for Multiscale Modelling

In [5, 6], the TANN framework has been used to homogenize the constitutive behavior of a micro-structured heterogeneous inelastic cell.

TANN are based on the thermodynamics theory of Internal State Variables, see [12]. The theory seeks to describe the state of a history-dependent material using a set of variables able to truck the microscopic irreversibilities occurring in the material, so as to permit a description of the state local in time. The material model is obtained by the definition of the Helmholtz free energy density function and the ISV evolution law. Micro-structured heterogeneous materials lack a straightforward definition of macroscopic ISVs. The Authors proposed discovering an a priori unknown set of ISVs, $\mathcal{Z}$, from dimensionality reduction of microscopic state information, encoded in what they refer to as a set of Internal Coordinates (IC), ξ. After defining the macroscopic state space,

$\mathcal{S}$, the TANN framework can be used to learn from data the homogenized behavior of the heterogeneous material in a thermodynamically consistent manner.

The training of the Helmholtz energy network, NN_ψ, is central to the thermodynamic compatibility ensured by TANN. The NN_ψ takes as input the state of the material to output the Helmholtz energy. The latter is automatically differentiated to return the stresses, $\boldsymbol{\Sigma} = \frac{\partial \mathrm{NN}_\Psi}{\partial \boldsymbol{E}}$, (fulfilling the First Principle) and to compute, together with the rates of ISV, $\dot{\mathcal{Z}}$, the rate of energy dissipation, $D = -\frac{\partial \mathrm{NN}_\Psi}{\partial \mathbf{Z}} \cdot \dot{\mathcal{Z}}$. This latter is constrained to be non-negative (fulfilling the Second Principle) using a regularization term included in the definition of the loss function, $\mathcal{L}$, to be minimized during the optimization procedure:

$$\mathcal{L} = \lambda_\Sigma \ell^\Sigma + \lambda_\mathrm{R}^\mathrm{D} \ell^\mathrm{D} \tag{2}$$

with λ_Σ and λ_R^D being weights for regulating the relative magnitudes of outputs, $\ell^\Sigma = \frac{1}{N}\sum_i \left\| \Sigma_\mathrm{i} - \frac{\partial NN_\Psi}{\partial E}_i \right\|_1$ and $\ell_\mathrm{R}^\mathrm{D} = \frac{1}{N}\sum_i \left\| \mathrm{Relu}\left\{ -\left(-\frac{\partial NN_\Psi}{\partial \mathcal{Z}} \dot{\mathcal{Z}} \right)_i \right\} \right\|_1$. The Re-ctified Linear Unit – Relu is defined as $Relu\{x\} = \{x, x > 0; 0, else\}$, $\|\cdot\|_1$ is used for the L_1 norm, N is the number of considered samples.

3 Material Model and Numerical Database

Numerical homogenization is frequently based on the statistical concept of a Representative Volume Element (RVE). The RVE is a region large enough to contain a sufficient number of statistically independent realizations of the field, yet small enough to be computationally tractable. The correlation length is used to define the spatial correlation of the field, but it doesn't dictate the size of the RVE. The RVE size may need to be larger than the correlation length if the latter is small, in order to capture enough independent realizations of the field. In the application at hand, we considered a simplified case, in which the selected RVE is a cube of unitary dimensions. The correlation length of the field has been assumed to be 0.3 times the RVE size. To generate the correlated random field, we used Markov covariance function, with x_{ij} being the lag distance matrix between two points of the domain and θ_c the correlation length.

$$\rho(x_{ij}, \theta_c) = \exp\left(-\frac{x_{ij}}{\theta_c} \right) \tag{3}$$

The random generation algorithm outputs a standard normal correlated field. The latter can be transformed to match any normal or log-normal distribution. We utilized the procedure to assign constitutive parameters to the geometric field derived from the coordinates of Gauss points of the computational model's elements. Table 1 reports the mean and the standard deviation of the parameters used.

Table 1. Mean ($\tilde{\mu}$) and standard deviation ($\tilde{\sigma}$) of the used constitutive parameters for the elastic-strain hardening plastic Drucker-Prager constitutive model with cap. Parameters without std. Dev. Are considered as homogeneous.

	E(Mpa)	ν(-)	d(kPa)	β(°)	α(-)	R(-)	K(-)	p'_0(kPa)
$\tilde{\mu}$	180	0.3	15	38	0.05	1.2	0.8	100
$\tilde{\sigma}$	1	0.01	0.5	0.1	-	-	-	0.5

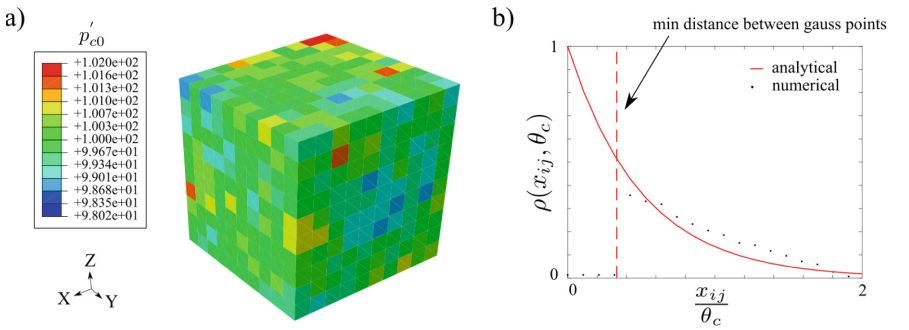

Fig. 1. a) Computational model of the RVE. The initial field of preconsolidation pressure p'_0 is depicted. b) Comparison of the the reconstructed numerical correlation function VS the analytical one.

The description of the constitutive parameters may be found at the link in reference [13], in the ABAQUS user's manual. Figure 1 depicts the computational model and the field of initial pre-consolidation pressure, the model has been initialized with.

For the training of TANNs a dataset of 25000 sample has been generated from 5 dataset of 1000 samples, to which a change of observer has been applied four times. This procedure has been implemented so that the network could learn objectivity from data. The RVE has been subjected to macroscopic random strain paths, after an initial monotonic volumetric compression up to 1e-3.

4 Results

After constructing the numerical database, we trained TANNs. At first, it was necessary to collect the microscopic data in the set of IC, ξ. Elastic and plastic deformations and maximum volumetric plastic deformations were utilized, the latter able to track volumetric hardening. The resulting number of IC's DoFs was 13000. POD was applied for the purpose of obtain ISVs, $\mathcal{Z}$, from the IC set. Figure 2 displays the normalized singular values and their cumulative sum for the first 200 POD modes. The sum approaches one with the considered modes, indicating a high degree of representativeness and a consequent small reconstruction error. The use of POD gives 0.0153 compression ratio. Starting from the reduced field, we applied further DRTs to obtain an extremely reduced

set of ISVs. The ability to reconstruct the high-dimensional IC field determined the choice of DRT to use.

According to the Coleman and Gurtin's theory, the set of ISVs must trace and represent the material's microscopic irreversible processes. Estimating the reconstruction error is necessary to assess the ISV's representativity. The latter is, however, difficult to compute. In general, it is not possible to define an analytic inverse function for nonlinear DRTs. In these cases, the reconstruction of the original field is accomplished by fitting an additional function onto data. If the reconstruction error, obtained by the composition of the reduction and reconstruction functions is small, it is possible to conclude that the composition approximates well the identity operator. However, the reduced set is representative only in a sense dictated by the reconstruction function; nothing ensures its representativeness in a thermodynamic sense. To achieve the latter property, the reduction function's fitting should be incorporated into the TANN training procedure. Indeed, the thermodynamic representativeness of the reduced set is so guaranteed by the successful training of the networks.

Following this reasoning, ICA, KPCA, and AE were considered. In ICA, there is an inverse function, so the reconstruction error can be calculated analytically. For KPCA, an analytical inverse doesn't exist. Nevertheless, after defining the kernel matrix, the application of PCA and the hierarchical sorting ensures that the reduced field is as representative as possible of the initial one, therefore, also thermodynamically representativeness is ensured. Finally, encoders do not have analytical inverse, but are simple enough to be incorporated into the TANN training procedure. In this case, it is not strictly necessary to define a decoder in order to achieve field reconstruction. Indeed, if training succeeds, it is concluded that the obtained reduced set is representative, no matter its reconstruction. However, if there is the need of reconstructing the microscopic field from the macroscopic one, it is possible to train an additional decoder, outside (or inside) the TANN training procedure.

A reduction dimension equal to 15 was chosen for all the DRTs to compare. This was chosen after ensuring a corresponding ICA's mean absolute reconstruction error of 1e-5. The training was done considering 10000 epochs, mini-batches of 1000 samples and Nesterov accelerated Adam's optimizer with learning rate 5e-5. The training was successful for all DRTs. In the case of ICA and KPCA, where the dimensionality reduction was performed outside of the TANN training, each training epoch required 3 ms, whereas the coupled training of the encoder required 15 ms per epoch, on a machine with 32 cores. The encoder had 3 hidden layers of decreasing size (150, 100, 50 neurons) and one output layer of 15 neurons, with tanh activation function. ICA fitting required 0.7 s. KPCA took 13 min and fifty seconds, including the fitting of an additional reconstruction function.

Trained TANNs were evaluated in inference mode on an unseen dataset. Figure 3 depicts the results of the prediction on a 3D strain-controlled path to which a volumetric strain was initially applied, followed by random increments. A very good agreement is achieved. The encoder's output contains some outliers. This is attributed to the coupling of the encoder's training with TANN, which causes a delay in achieving learning convergence in 10000 epochs, resulting in a decreased accuracy. With ICA and KPCA, the asymptotic learning value is found very early, between 300 and 200 epochs.

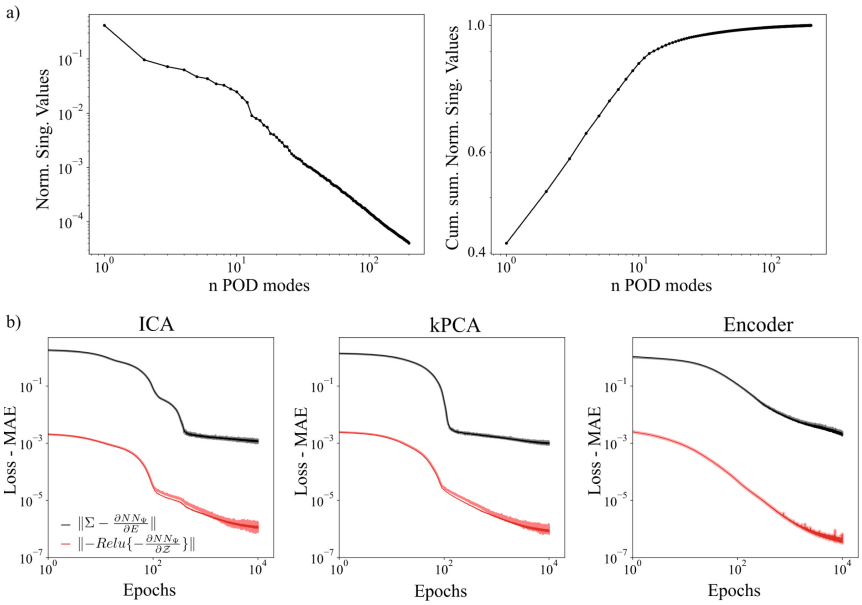

Fig. 2. a) Normalized singular values and their cumulative sum considering the first 200 POD modes obtained from the set of IC, ξ. b) Learning curves expressed in terms of the Mean Absolute Error – MAE, considering ICA, kPCA and the Encoder dimensionality technique. 10000 epochs have been used. Solid lines represent results on the training set, shaded lines below the solid ones on the validation set.

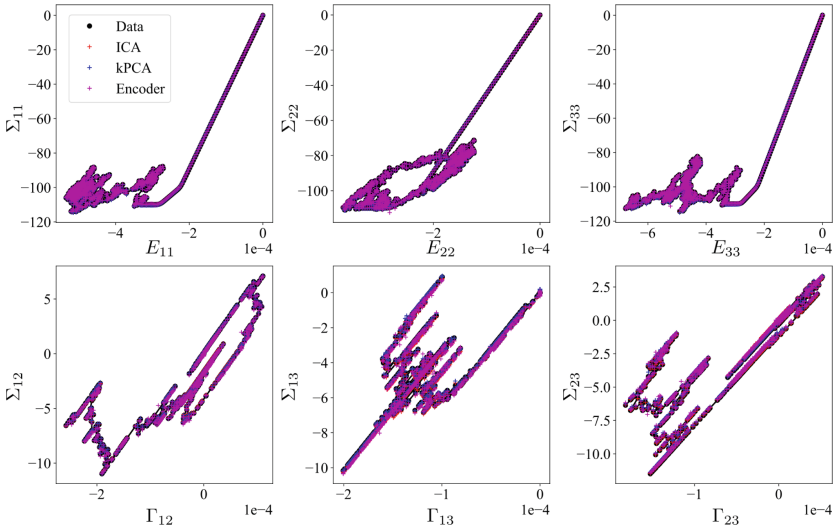

Fig. 3. TANN predictions in inference mode on a set of unseen data. The network's predictions obtained considering the set of ISV obtained with ICA, kPCA and the encoder are compared.

5 Conclusions

In this study, three DRTs have been applied on top of POD to define a set of ISVs for the macroscopic state space definition of a heterogeneous medium. TANNs were subsequently trained on a dataset of data derived from numerical simulations. The RVE was obtained by assigning spatially correlated constitutive parameters to the elements of the model's structured mesh. All DRTs produced satisfactory results. ICA returned excellent results while requiring the least amount of computational time, ensuring the possibility of reconstructing microscopic fields. KPCA demonstrated to be a valid alternative, attractive if a linear method, like ICA, fails to generate satisfactory results. Unlike encoders, KPCA fitting can be decoupled from TANN training. This is time saving and makes it easier to fine-tune neural networks in the TANN architecture.

References

1. Wang, K., Sun, W.: A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning. Comput. Methods Appl. Mech. Eng. **334**, 337–380 (2018). https://doi.org/10.1016/j.cma.2018.01.036
2. Hall, E.J., Taverniers, S., Katsoulakis, M.A., Tartakovsky, D.M.: GINNs: graph-informed neural networks for multiscale physics. J. Comput. Phys. **433**, 110192 (2021). https://doi.org/10.1016/j.jcp.2021.110192
3. Sorini, A., Pineda, E.J., Stuckner, J., Gustafson, P.A.: A convolutional neural network for multiscale modeling of composite materials. In: AIAA Scitech 2021 Forum, p. 0310 (2021)
4. Masi, F., Stefanou, I.: Multiscale modeling of inelastic materials with Thermodynamics-based Artificial Neural Networks (TANN). Comput. Methods Appl. Mech. Eng. **398**, 115190 (2022)
5. Masi, F., Stefanou, I.: Evolution TANN and the discovery of the internal variables and evolution equations in solid mechanics. arXiv preprint arXiv:2209.13269 (2022)
6. Piunno, G., Masi, F., Stefanou, I., Jommi, C.: Multi-scale modelling of natural composites via thermodynamics-based artificial neural networks. In: Congrées Français de Mècanique – CFM, AFM (2022)
7. Lumley, J.L.: The structure of inhomogeneous turbulent flows. Atmos. Turbul. Radio Wave Propag. 166–178 (1967)
8. Géron, A.: Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, Inc. (2019)
9. Kutz, J.N.: Data-Driven Modeling & Scientific Computation: Methods for Complex Systems & Big Data. Oxford University Press, Oxford (2013)
10. Schölkopf, B., Smola, A., Müller, K.R.: Kernel principal component analysis. In: Gerstner, W., Germond, A., Hasler, M., Nicoud, JD. (eds.) Artificial Neural Networks (ICANN 1997). LNCS, vol. 1327, pp. 583–588. Springer, Heidelberg (1997). https://doi.org/10.1007/BFb0020217
11. Myers, D.E.: Vector conditional simulation. In: Armstrong, M. (ed.) Geostatistics. Quantitative Geology and Geostatistics, vol. 4, pp. 283–293. Springer, Dordrecht (1989). https://doi.org/10.1007/978-94-015-6844-9_21
12. Coleman, B.D., Gurtin, M.E.: Thermodynamics with internal state variables. J. Chem. Phys. **47**, 597 (1967)
13. ABAQUS: https://classes.engineering.wustl.edu/2009/spring/mase5513/abaqus/docs/v6.6/books/usb/default.htm?startat=pt05ch18s03abm29.html